DOCUMENT RESUME

ED 095 792 HE 005 868

AUTHOR	Kinsel, John B.; Latham, Jefferson
TITLE	A Comparison of Two course Evaluation Forms: The Hofstra University Course and Teacher Evaluation Questionnaire and the New College Course Evaluation Questionnaire. Report No. 3.
INSTITUTION	Hofstra Univ., Hempstead, N.Y. New Coll. Educational Research Office.
PUB DATE	Jun 74
NOTE	15p.; Some lines of text may not reproduce clearly
EDRS PRICE	MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS	*Course Evaluation; Evaluation; *Higher Education; *Measurement Instruments; *Questionnaires; *Teacher Evaluation
IDENTIFIERS	Hofstra University; *New College

ABSTRACT

This report concerns a comparison of the course and teacher evaluation form used by New College of Hofstra with another form recently developed for use in other units of Hofstra University. The comparative analysis is based on data derived from administration of the Hofstra University Course and Teacher Evaluation Questionnaire and the New College Evaluation Questionnaire to 368 New College respondents at the conclusion of the spring semester 1974. The purpose of this study was to compare results on the two forms to determine: (1) if similar information was gained from the two forms, (2) if one or the other of the forms alone provided information appropriate to the needs of New College, and (3) if in the future it would be necessary to administer both forms. Overall, it might be concluded that the two forms do get at basically the same information, since significant correlations do exist between many items and many items do have "face" or content similarity. Yet, as previously indicated, much higher correlations might have been expected among certain items that are in fact identical. The possibility of some systematic bias in responding, as suggested by the review of individual respondee sheets, and by the relatively low intercorrelations between forms with relatively high interacorrelations between items on the same form, limits any extensive interpretation of comparability of forms. Reproduced from best available copy. (Author/PG)

ED 095792

A COMPARISON OF TWO COURSE EVALUATION FORMS: THE HOFSTRA UNIVERSITY COURSE AND TEACHER EVALUATION QUESTIONAIRE AND THE NEW COLLEGE COURSE EVALUATION QUESTIONAIRE

John B. Kinsel
Coordinator, New College Educational
Research Office

and

Jefferson Latham
Research Associate, New College
Educational Research Office

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BEST COPY AVAILABLE

June, 1974
NCERD Report #3

HE 005 868

A Comparison of Course Evaluation Forms

INTRODUCTION

This report concerns a comparison of the course and teacher evaluation form used by New College of Hofstra with another form recently developed for use in other units of Hofstra University. The comparative analysis is based on data derived from administration of the Hofstra University Course and Teacher Evaluation Questionaire (hereafter referred to as Form H) and the New College Course Evaluation Questionaire (hereafter referred to as Form N) to 368 New College respondents at the conclusion of the Spring semester, 1974.

The current edition of the New College questionaire (Form N) has been used for evaluation of courses by students for the past year. Similar questionaires were used in the previous academic year, 1972-73 and, in general, New College has experimented with a variety of different forms since 1969. The Hofstra University questionaire, (Form H) was developed during the current academic year and was administered voluntarily in a number of courses at the end of the Fall semester, 1973-74. Form H was then to be administered in all undergraduate courses at Hofstra University during the Spring semester.

Each of the evaluation forms has been designed with specific purposes or audiences in mind. Form H was developed in response to student requests for an "objective" evaluation of instructors and courses. At its March 15, 1974, meeting, the University Faculty voted to make administration of the questionaire mandatory in all undergraduate courses. This motion and decision were made with particular reference to Faculty Policy Series #15 which states, in part, that student evaluations and student opinions may be used in making tenure decisions for Faculty. Consequently, the use of Form H appears to serve two purposes: first, to provide information to students regarding faculty and courses, and, secondly, to provide a uniform system for collecting student opinions to be used by appropriate parties in making personnel decisions regarding individual faculty.[1]

Form N represents the current means by which New College implements its long standing practice of faculty and course evaluation. The New College faculty and administration have agreed on the following uses of evaluation results:

---

1 It is interesting to note that while administration of the questionaire (Form H) is now mandatory in all undergraduate courses, the results need not necessarily be used in tenure decisions. Faculty Policy Series #15, IV, a. (b) states, "any opinion the candidate may wish to present for inclusion in the file of the proceedings of the tenure committee to support his candidacy, such as student evaluations . . ." (underscoring ours)

(a) to provide information and feedback to the faculty member about perceptions of his course by students, (b) to provide the New College Dean's Office and appropriate faculty with an evaluation of the effectiveness of the various programs within the College, and (c) to provide information about contributions of new and/or experimental [illegible] individual programs. Administration of the New College form is handled by the New College Educational Research Office and is mandatory for all courses. The Dean and faculty of New College, furthermore, agree that information from courses and faculty evaluations will not be used in personnel decisions unless the faculty person voluntarily and at his/her initiative submits the material for such use.

PURPOSE OF THIS STUDY

Because of the mandatory status of the University questionnaire (form H) and for continuity of program evaluation at New College, New College administered both Forms H and N at the conclusion of the Spring semester. The purpose of this study was to compare results on the two forms to determine (1) if similar information was gained from the two forms, (2) if one or the other of the forms alone provided information appropriate to the needs of New College, and (3) if, in the future, it would be necessary to administer both forms.

METHODOLOGY

Subjects and administration

Both Forms H and N were administered to 368 respondents in 34 classes at New College. Students completed the two questionaires on the same Opscan answer sheet so that pairwise comparisons could be made. Most respondents were familar with Form N since they had completed this form on three previous occasions during the year. No special instructions were given with reference to the completion of Form H with the exception that the students were cautioned several times to read the University Form instructions carefully and to be aware of the direction of the scaling.[2]

Data analysis

The first step in the data analysis was to identify pairs of questions on the two forms which were similar. This was accomplished in two ways. First, the investigators reviewed the questions on each form in terms of their content and apparent "face" similarity. Then, Pearson Product Moment correlations were calculated for each pair of questions. (The resultant pairing of questions as listed in Table I are the final result of these procedures taken together.)

To determine the dimensions involved in the questionaires, a factor analysis was performed on each questionaire. This procedure is completed

---

2 The problem of assumed reversal of scaling will be discussed later in this report.

regularly after each administration of the New College Form.

RESULTS AND DISCUSSION

Comparability of Information

A pairwise comparison was performed among the 22 questions on the University Form and the 18 evaluative questions on the New College Form. Because of the method of scaling (where 3 is a "perfect" score instead of 5) questions 17-19 on Form H are not discussed in the analysis. Also because the scaling directions are reversed on the forms, all the resulting correlations between forms are negative but can be interpreted as if they were positive.

To provide a frame of reference for interpreting correlation coefficients between forms, intracorrelations of items on the same form were also reviewed. It was noted that there were an unexpectedly high number of high correlations among the questions on Form H. Omitting questions 17-19 because of scaling problems, nineteen questions remained. Choosing an arbitrary correlation level of .70, there were 32 pairs of questions on Form H which exceeded that level. By comparison, Form N had only 17 pairs which exceeded .70.

On the other hand, few high correlations between questions on the different forms were found. The highest correlation was -.51 between "Overall rating of the instructor inside the classroom" (Form N) and "Would you take another course with this instructor?" (Form H). The majority of coefficeints centered around -.30 with a range of -.01 to -.51. Because of the sample size, almost all correlations above -.20 were significant at the .001 level.

Table I provides a list of potentially comparable items on the two forms along with the resulting correlation coefficients between the items. This list was initially determined on the basis of "face" similarity of items. Then, any items which correlated at -.35 or higher were added. As indicated within the Table, some items on Form N (3, 4, 13, and 16) had no comparable questions on Form H and no correlations as high as -.35. Similarly, the same is true of items on Form H (10, 11, 16, 18, and 22). Items 7 ("How valuable was the exam(s) to you as a learning tool?") and 11 ("Was and approporiate level of prior knowledge assumed?") on Form N had comparable items on Form H but could not be evaluated because of the different scaling systems used for the responses.

TABLE I

Comparison of questions on Form N and Form H

Question # Form N	Form H	r*
1	1	-.38
	2	-.41
	3	-.20
	8	-.42
	9	-.44
	15	-.40
2	4	-.33
	5	-.32
3	no comparable question	
4	no comparable question	
5	12	-.29
6	12	-.29
7	17	no comparable data because of scaling difference
8	1	-.42
	2	-.45
	3	-.18
	7	-.25
	8	-.42
	9	-.51
	10	-.28
	11	-.34
9	6	-.45
10	1	-.43
	2	-.45
	3	-.17
	8	-.44
	15	-.43
11	19	no comparable data becaus of scaling difference
12	13	-.37
	15	-.47
	8	-.44
13	15	-.39
14	14	-.30

TABLE I
continued

Question # FORM N	FORM H	r *
15	20	-.39
16	no comparable item	
17	1	-.42
	2	-.41
	4	-.32
	8	-.37
	9	-.50
	13	-.37
	14	-.39
	15	-.44
18	1	-.39
	2	-.39
	7	-.37
	8	-.46
	13	-.38
	14	-.35
	15	-.45

* $p < .001$ for all r's reported.

Given the "face" similarity between many of the questions, the resultant correlations are lower than might have been expected. On one question which was virtually identical on both forms - "Availability of instructor outside the classroom" (Form N) compared to "Outside the class the instructor was: available to unavailable" (Form H) - the correlation was -.41. Since the same students were measuring the same instructor on a virtually identical dimension at the same point in time, the resultant correlation might have been expected to more closely approximate 1.00. The fact that it did not suggested that some systematic bias may have been operating.

A review of individual response sheets of some students revealed that in several cases students had rated one item "most positively" (e.g. "A" on form N), and the similar item as "most negative" on the other form. Since it is highly unlikely that a student would intentionally reverse his ratings in such a way, the most probable explanation has to do with the directionality of the scaling. The scaled values on Form H run from 5 (most positive) to 1 (most negative). These values run horizontally from left to right, whereas the values on the Opscan sheet run vertically from 1 to 5. Furthermore, there may be, in the minds of some students, a "psychological" tendency to think in terms of using 1 as the highest or most positive value (i.e. "first"). Consequently, from a human engineering point of view, these two factors may have combined resulting in unintentional confusion on the part of the student as he responded. If so, then, the reliability of the results of Form H must be seriously questioned.[3]

Summary: Overall, it might be concluded that the two forms do get at basically the same information since significant correlations do exist between many items and many items do have "face" or content similarity. Yet, as previously indicated, much higher correlations might have been expected among certain items which are in fact identical.

The possibility of some systematic bias in responding as suggested by the review of individual response sheets and by the relatively low inter-correlations between forms with relatively high intracorrelations between items on the same form, limits any extensive interpretation of comparability of forms.

Factor Analysis

Theoretically, both student evaluation questionaires are designed to gather data on more than one dimension. Each question should, as far as possible, require a simple rather than a compound answer. Ideally, singularity of demensions is sought. Therefore, questions which measure each dimension and correlate highly are used to represent a dimension under investigation. The result is that a 19 or 20 item questionaire will measure three or hopefully more underlying dimensions. With this theoretical model in mind, a factor analysis was performed on each of the questionaires.

3 In development of the current New College questionaire, an attempt was made to avoid the problems of scaling directionality and "psychological set" by asking the students to rate each question from A to E with A being "very good" and E "very poor".

The analysis of Form N produced two factors which were labelled (1) "Instructor inside the classroom" and (2) "Students own effort as reflected in attendance, study time and involvement." To put this result into perspective, in factor analysis of the three previous administrations of Form N, three factors consistently emerged: (1) Instructor inside the classroom, (2) External parameters of the class such as assignments and projects, and (3) student involvement. For some reason, the first two factors merged into one in this current evaluation. Several types of rotation were performed in an attempt to restore the 3 dimensionality of the evaluation. However all methods produced a clear two factor structure. By interpretation, from the clustering on factor I, it is possible to separate factor I into the two original factors. This method indicates that both factors are of the same dimensionality, but that the instructor inside the classroom is a more powerful cluster than the parameters of the course.

The factor analysis of Form H yeilded only one factor which accounted for 66% of the variance. No variable loaded less than .61 on this factor. For want of a better name, this factor might be labelled "Did you like the instructor?" Again, by interpretation of the clustering on the factor, two additional dimensions can be hypothesized. One might be identified as the (1) instructor outside the class along with his tolerance for disagreement and the other as the (2) students own efforts.

In general, the factor analysis of Form H suggests that the questions are probably multi-dimensional and that all questions may be calling for an evaluation of one underlying dimension.

It must be remembered, that the factor analysis of Form H discussed above was done with New College students only. It would be informative to know if the same pattern holds for the University as a whole.

CONCLUSIONS AND RECOMMENDATIONS

(1) It is recommended that New College continue to use its own form, and that with the consent of the New College faculty, make the results of individual courses available for publication assuming the University continues the policy of making public the course evaluation results. Reasons for this recommendation include the following: (a) Comparison of questions between forms suggests that there exists "face" similarity between questions and significant correlations between questions. (b) The factor analysis suggests the single dimesion emerging on Form H is roughly equivalent to the primary factor emerging on Form N, namely the "Instructor". Other reasons not mentioned so far in this analysis include: (c) the greater familiarity which the New College faculty and students have with the current form and (d) maintainance of continuity in program evaluation for New College.

(2) Modifications might be made in the New College form to accomadate items which are included in the University form but not on the New College Form. These include questions about the instructor's knowledge of graduate and career opportunities in his field, the value of laboratory experiences (where appropriate), fairness of grading, and the contribution of other students. Similarly, the University might consider adopting questions from the New College from such as the learning value of assignments and exams, helpfulness in writing skills, and "overall" ratings.

Recommendations regarding Form H

Based upon the analysis of Form H, our observations make while administering the University Form, and our own experience in developing and using course evaluation forms, the following suggestions are made regarding the format of the current University form:

(1) The scaling system might be reversed to made it consistent with the Opscan sheet and with the normal/typical "psychological" directionality of the respondent.

(2) All question should use the same scaling system. The current form requires that the student change his "response set" for questions 17-19, and then revert back. This combined with problems of directionality may cause the student added difficulty in responding.

(3) The instructions are long and complicated and may contribute to fatigue and a certain "negativism" before the student even starts to answer the questions. Similarly, there is no indication at the outset as to what the purpose of the evaluation is.

(4) The way the form is set up is somewhat awkward. That is, the positive attribute is separated by five or six spaces from the #5 (or positive response) and the negative attribute is directly adjacent to the #1. Whether this is a significant factor is difficult to judge, but it does have the effect of visually distorting the scale.

(5) The narrative comments are separated from the objective questions. Our experience at New College has been that the narrative comments are often the most informative and valuable to the instructor. Similarly, we have found that narrative comments are most readily illicited when space is provided immediately after the "objective" questions, i.e. narrative comments about the instructor immediately following the questions about the instructor.

General Comments

The preceeding points are important with regard to the technical format of the questionaire, and suggest changes which may enhance the usability and readibility of the form. However, of much greater concern is the overall reliability and validity of this or any such measure. This issue is of particular importance if the information is to be used by students in making choices of courses and/or if it is to be used by others in making important personnel decisions.

Recently, Colleges and Universities across the country have been adopting, abondoning, and experimenting with course and teacher evaluation of all types and varieties. In general, research has not kept pace with these developments (Villano, Rosenstock, and Estes, 1974). Many schools have abandoned the use of evaluations because of their lack of validity or meaningfulness and the findings reported in the literature are often characterized as inconsistent and contradictory (Costin, Greenough, and Menges, 1971).

General Comments

Our limited analysis suggests that the University form, as it now exists, may not be effectively discriminating different characteristics or dimensions of the course and the instructor. If the results are to be used as defined, then it would appear that more than the "likability" of a given faculty member should be taken into consideration. In general, if Hofstra is going to institute a mandatory system of course evaluations, a careful evaluation of the meaningfulness and accuracy of information resulting is indeed essential.

References

Costin, F., Greenough, W., & Menges, R. Student ratings of college teaching: realiability, validity, and usefulness. Review of Educational Research, 1971, 41(5), 511-535.

Villano, M.W., Rosenstock, E.H. & Estes, C. A decade with a student course evaluation form at a major university. Paper presented at AERA Annual Meeting, Chicago, April, 1974.

NEW COLLEGE, HOFSTRA UNIVERSITY

Course Evaluation Questionnaire

Course _____ Instructor _____

Session _____ Month/Year _____ Your Academic Year _____

Your Concentration_____ Is this a required course?_____ Grade you expect____

This questionnaire is designed to evaluate and improve New College courses. All answers will be anonymous.

Please rate each of the dimensions of this course according to the following code:

A = Very Good/Well B = Good/Well C = Average
D = Poor E = Very Poor Blank = Not Applicable

1. _____ Quality of class presentations, lectures and demonstrations.

2. _____ How well did class discussion help in the learning experience of this class?

3. _____ How valuable were the assigned readings to you in your studies of this course topic?

4. _____ Value of papers, projects, problems and other additional assignments.

5. _____ Did this course have a major focus? Did it have a direction which you understand?

6. _____ To the extent that you agreed that this course had a major focus, was that focus or objective meaningful to you?

7. _____ How valuable was the exam or exams as a tool to help you to learn?

Comments: Use this space to add further comments you want to make concerning the first seven questions which are on specific aspects of the course.

_____

_____

_____

8. _____ Overall rating of the instructor inside the classroom.

9. _____ Availability of the instructor outside of the classroom.

10. _____ How well did the professor present the material/content of the course in a logical order that helped you to learn.

11. _____ Did the course assume an appropriate level of prior knowledge and/or experience by the students?

Comments about the Instructor & Instruction: _____

_____

_____

12. _____ What was the extent of learning about the field of inquiry.

13. _____ What was the extent of learning about skillful writing.

14. _____ To what extent did you become involved in the course?

15. _____ To what extent did you expend time and effort in studying for this course?

16. _____ What was your attendance in class?

Comments on your Personal Involvement: _____

_____

_____

17. _____ Overall rating of the course.

18. _____ To what extent were your expectations in taking the course achieved?

Mark A for each of the following items which were a reason for your taking the course. Leave blank those items which were not reasons for your taking the course..

19. _____ Needed to meet a requirement.

20. _____ Needed to prepare for a career or graduate school.

21. _____ The subject of the course was interesting.

22. _____ The professor had a good reputation.

23. _____ The course was less demanding required less work.

24. _____ This was the only course which fit into my schedule.

25. _____ All other reasons, please specify on the printed sheet.

General Comments: _____

_____

_____

_____

HOFSTRA UNIVERSITY

COURSE AND TEACHER EVALUATION

Questionnaire

Directions: Here is a list of attributes concerning either your instructor and/or your course. Read each attribute and its rating scale. When you decide on a rating, blacken the appropriate space on your answer sheet (Questions 1-22). Also fill in whether you are part or full-time (day or evening), and your student status (freshman, sophomore, etc.). In the student number box use the first three columns (left most columns) to enter your cumulative average--do not include a decimal point. In the sex box, blacken "M" if this course was required for major or for degree; and the "F" if it is not required. Use the pencil supplied and erase fully if you wish to change your answer.

Remember, at the top of each part are definitions of the scores; these differ for each part. Not every item is appropriate to each course; inappropriate items should be omitted.

Part I - The Instructor	Score Definition 5= Perfect score 3= Average score 1= Lowest Possible Score
1. The instructor's organization of subject matter was (Clear)	5 4 3 2 1 (Confusing)
2. The organization of each class was (Well organized)	5 4 3 2 1 (Unorganized)
3. The instructor's knowledge of the subject matter was (Outstanding)	5 4 3 2 1 (Inadequate)
4. Did the instructor see to it that the classroom situation was conducive to discussion? (Very much)	5 4 3 2 1 (Not at all)
5. Did the instructor see to it that the classroom situation was conducive to questioning? (Very much)	5 4 3 2 1 (Not at all)
6. Outside of class the instructor was (Available)	5 4 3 2 1 (Unavailable)
7. The instructor's attitude toward disagreement was (Tolerant)	5 4 3 2 1 (Intolerant)
8. Intellectually the instructor was (Stimulating)	5 4 3 2 1 (Not at all)

9. Would you take another course with this instructor? (Definitely) 5 4 3 2 1 (Never)

10. How much knowledge did the instructor have of graduate programs in his field. (A Great Deal) 5 4 3 2 1 (None)

11. How much knowledge did the instructor have of career opportunities in his field? (A Great Deal) 5 4 3 2 1 (None)

Part II - Course

Score Definition
5= Perfect Score
3= Average Score
1= Lowest Possible Score

12. Did the course cover the material it was supposed to: (Completely) 5 4 3 2 1 (Not at all)

13. Has your ability to evaluate evidence in this field increased? (A Great Deal) 5 4 3 2 1 (Not at all)

14. As a result of this course your interest in the subject matter has (Increased) 5 4 3 2 1 (Decreased)

15. The amount learned from this course was (A Great Deal) 5 4 3 2 1 (Nothing)

16. How much did the laboratory part of this course contribute to your knowledge of the subject? (A Great Deal) 5 4 3 2 1 (Not at all)

Part IIB - Course

Score Definition
3= Perfect Score

17. The quality of tests was (Too Difficult) 5 4 3 2 1 (Too easy)

18. The grading was (Too hard) 5 4 3 2 1 (Too easy)

19. The level of course was (Too difficult) 5 4 3 2 1 (Too easy)

Part III - The Student

Score Definition
5= Perfect Score
3= Average
1= Lowest Possible Score

20. The amount of effort you put into this course was (A Great Deal) 5 4 3 2 1 (Nothing)

21. Were you prepared for the level on which this course began? (Entirely) 5 4 3 2 1 (Not at all)

22. Did the other students in the class help the classroom exchange be constructive (Very much) 5 4 3 2 1 (Not at all)

Part IV - Open Response Questions (Optional)

Instructor_____

Course_____

1. The following strengths impressed me most about this course/instructor:

2. The following weaknesses were evident to me about this course/instructor:

3. If the course had a lab or discussion section, please give any comments on how they helped, or how they can be improved:

4. Please make any other comments about this course, the instructor, or this evaluation questionnaire: